Markov Chain
Basic Concepts

Laura Ricci

Dipartimento di Informatica

24 luglio 2012

PhD in Computer Science

P2P System Evolution: Brief Overview

- *Unstructured Approaches*
  - random connections
  - search mechanisms: TTL-enhanced flooding
  - examples: Gnutella, Kazaa, BitTorrent,...
  - *pro:* simplicity, *cons:* limited scalability
- *Structured Approaches*
  - Distributed Hash Tables, Delaunay Overlay
  - examples: Kademlia, implementation in eMule/Bittorrent
  - *pro*: routing efficiency, *cons*: consistency mainteinance in presence of churn
- *Stochastic Approaches*
  - Gossip, Random Walk
  - simple, based on a well known mathematical theory
  - exploited also for distributed crawling of social networks

P2P System: Stochastic Approaches

- unstructured/structured approaches presented in the Course "P2P Systems" of the Master Degree in Computer Science
- here we focus on stochastics appoches
- mathematical background
  - Markov Chains, Random Walk
- applications:
  - Distributed sampling for extimating network parameters
  - Epidemic information diffusion
  - Distributed computation af aggregate values

Outline

Markov Chain: Basic Characteristics

- Markov Chain: describes a system whose states change *over time*
  - discrete time stochastic process
- Changes are governed by a *probability distribution*.
- The next state only depends upon the current system state
  - the path to the present state is not relevant
- Class of random process useful in different areas
  - developments in theory and applications in recent decades.

Markov Chain Specification

A sequence of Random Variables $\{X_0, X_1, \ldots X_n\}$: X_i describes the state of the system at time i

To specify a Markov Chain we need:

- a finite or countable set of states S
  - $S = \{1, 2, \ldots, N\}$ for some finite N
  - the value of the random variables X_i are taken from S
- an initial distribution π_0, $\pi_0(i) = \mathbb{P}\{X_0 = i\}$: probability that the Markov Chain starts out in state i
- the probability transition rules

The Probability Matrix

- The probability matrix $P = (P_{ij})$ specifies the transition rules
- if the size of S is N, P is a $N \times N$ stochastic matrix
  - each entry is non negative
  - the sum of each row is 1
- P_{ij} is a conditional probability: defines the probability that the chain jumps to state j, at time n + 1, given that it is in state i at time n,

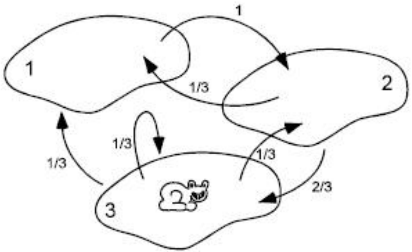
$$P_{ij} = P(X_{n+1} = j \mid X_n = i)$$

- We assume *Time Homogeneity*: the probability does not depend on the time n, but probability only on depends state i and j

Time Homogeneity

- "where I go next given that I am in state s at time x is equal to where I go next given that I am in the same state s at time $y \neq x$"
- every time the chain is in state s, the probability of jumping to another state is the same
- the same probability over time
- we will assume time homogeneity in the following

The Markov Frog



A frog hopping among lily pads

- State Space: $S = \{1, 2, 3\}$ represent the pads
- Initial Distribution: $\pi_0 = (1/2, 1/4, 1/4)$
- Probability Transition Matrix:

$$P = \begin{pmatrix} 0 & 1 & 0 \\ 1/3 & 0 & 2/3 \\ 1/3 & 1/3 & 1/3 \end{pmatrix}$$

  - defines the probabilities of jumping from one state to another one

The Markov Frog

- the frog chooses its initial position X_0 according to the initial distribution π_0
- to this purpose, the frog can ask its computer to generate a uniformly distributed random number U_0 in the interval $I = [0, 1]$ and then taking

$$X_0 = \begin{cases} 1 & \textit{if} \quad 0 \leq U_0 \leq 1/2 \\ 2 & \textit{if} \quad 1/2 < U_0 \leq 3/4 \\ 3 & \textit{if} \quad 3/4 < U_0 \leq 1 \end{cases}$$

- for instance, if $U_0 = 0.8419$ then $X_0 = 3$
  - the frog starts on the third lily pad

The Markov Frog

- Evolution of the Markov Chain: the frog chooses a lily pad to jump
    - state after the first jump = value of the random variable X_1
- the frog starts from lily pad 3 so look at the probability distribution in row 3 of P, namely (1/3,1/3,1/3)
- again, generate a uniformly distributed random number U_1 in the interval $I = [0, 1]$ then takes

$$X_1 = \begin{cases} 1 & \textit{if} \quad 0 \leq U_1 \leq 1/3 \\ 2 & \textit{if} \quad 1/3 < U_1 \leq 2/3 \\ 3 & \textit{if} \quad 2/3 < U_1 \leq 1 \end{cases}$$

- if $U_1 = 0.1234$, then $X_1 = 1$
    - the frog jumps from the lily pad 3 to lily pad 1
    - $X_1 = 1$ there is no choice for the value of X_2, it must be 2
    - and so on....

The Markov Property

- in the previous example

  $$\mathbb{P}\{X_3 = j \mid X_2 = 2, X_1 = 1, X_0 = 3\} = \mathbb{P}\{X_3 = j \mid X_2 = 2\} \quad \forall j$$

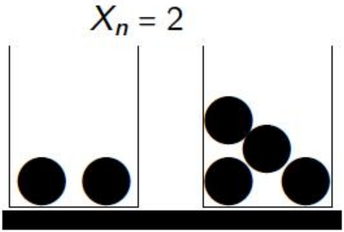
- the only information relevant to the distribution to X_3 is the information that $X_2 = 2$
- $X_0 = 3$ and $X_1 = 1$ may be ignored!

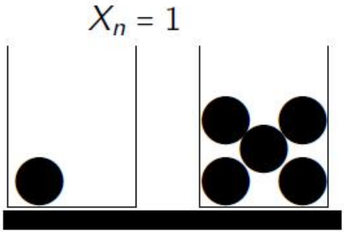
Definition

A stochastic process X_0, X_1,....satisfies the Markov Property if $\mathbb{P}\{X_{n+1} = i_{n+1} \mid X_n = i_n, X_{n-1} = i_{n-1}, \dots, X_0 = i_0\} = \mathbb{P}\{X_{n+1} = i_{n+1} \mid X_n = i_n\}$

Markov Chains: Ehrenfest Chain

There is a total of 6 balls in two urns, 4 in the first and 2 in the second. We pick one of the 6 balls at random and move it to the other urn.
X_n number of balls in the first urn, after the nth move.





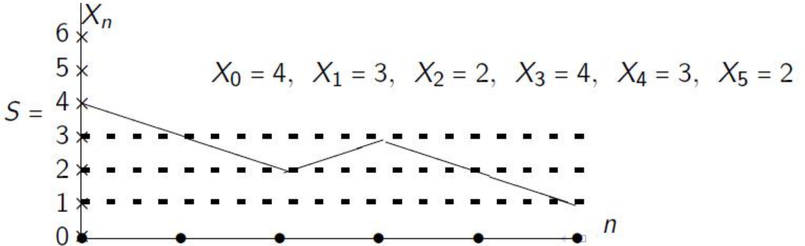
Markov Chains: Ehrenfest Chain

$$\mathbb{P}(X_0 = 4) = 1$$

$$\mathbb{P}(X_1 = j) = \begin{cases} 4/6 & j = 3 \\ 2/6 & j = 5 \\ 0 & otherwise \end{cases}$$

$$\mathbb{P}(X_{n+1} = l \mid X_n = j) = \begin{cases} j/6 & l = j - 1 \\ (6 - j)/6 & l = j + 1 \\ 0 & otherwise \end{cases}$$

In 5 unites of time $X_0, \ldots X_5$ might follow the following path:



Ehrenfast Chain: Probability Transition Matrix

$$P = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1/6 & 0 & 5/6 & 0 & 0 & 0 & 0 \\ 0 & 2/6 & 0 & 4/6 & 0 & 0 & 0 \\ 0 & 0 & 3/6 & 0 & 3/6 & 0 & 0 \\ 0 & 0 & 0 & 4/6 & 0 & 2/6 & 0 \\ 0 & 0 & 0 & 0 & 5/6 & 0 & 1/6 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

$$\begin{tikzcd} 0 \arrow[r, shift right, "1"'] & 1 \arrow[l, shift right, "1/6"'] \arrow[r, shift right, "5/6"'] & 2 \arrow[l, shift right, "2/6"'] \arrow[r, shift right, "4/6"'] & 3 \arrow[l, shift right, "3/6"'] \arrow[r, shift right, "3/6"'] & 4 \arrow[l, shift right, "4/6"'] \arrow[r, shift right, "4/6"'] & 5 \arrow[l, shift right, "5/6"'] \arrow[r, shift right, "5/6"'] & 6 \arrow[l, shift right, "1"'] \end{tikzcd}$$

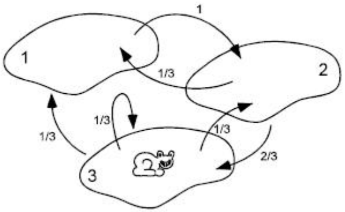
Markov Chains and Random Walks

- given a graph and a starting point
- select one of its neighbours at random, and move to this neighbor
- then select a neighbour of the new point point at random, and move to it, and so on ...
- the (random) sequence of points selected in this way is a *random walk* on the graph.
- there is a strict relation between Random Walk and Markov chains
  - time-reversible Markov chains can be viewed as random walks on undirected graphs
  - we will see this later ...

Random Walks and Complex Networks

- what is the connection between Random Walk and P2P/Social Networks?
- a simple way to analyse complex graph
  - gather information about all nodes and links at a single server
  - exploit classical graph theory
  - unfeasible solution due to the huge dimension of the graphs
- an alternative solution
  - compute a node sample through a random walk
  - exploit the node sample to estimate network properties

It's All Just Matrix Theory?



- what is the probability p_{ij}^n that, given the chain in state i, it will be in state j, n step after?
- if we start on lily pad 3, what is the probability of being on lily pad 1, after 2 steps?

  $$p_{31}^2 = \quad p_{31}p_{11}+p_{32}p_{21}+p_{33}p_{31} =$$

  $$\frac{1}{3}0 + \frac{1}{3}\frac{1}{3} + \frac{1}{3}\frac{1}{3}=\frac{1}{9}+\frac{1}{9}=\frac{2}{9}$$

- This is the dot product third row-first column of P
  - this returns the 1,3-entry of the product of P with itself.

It's All Just Matrix Theory?

The generalization of the previous observation leads to the following theorem.

Theorem

Let P be the transition matrix of a Markov chain. The ij-th entry of the matrix P^n gives the probability that the Markov chain, starting in state s_i, will be in state s_j after n steps.

The power of the transition matrix gives interesting information about the evolution of the process

It's All Just Matrix Theory?

$$P^1 = \begin{array}{c c} & \begin{array}{ccc} pad_1 & pad_2 & pad_3 \end{array} \\ \begin{array}{c} pad_1 \\ pad_2 \\ pad_3 \end{array} & \begin{pmatrix} 0 & 1 & 0 \\ 0.333 & 0 & 0.666 \\ 0.333 & 0.333 & 0.333 \end{pmatrix} \end{array}$$

$$P^2 = \begin{array}{c c} & \begin{array}{ccc} pad_1 & pad_2 & pad_3 \end{array} \\ \begin{array}{c} pad_1 \\ pad_2 \\ pad_3 \end{array} & \begin{pmatrix} 0.333 & 0 & 0.666 \\ 0.221 & 0.554 & 0.221 \\ 0.221 & 0.443 & 0.332 \end{pmatrix} \end{array}$$

$$P^3 = \begin{array}{c c} & \begin{array}{ccc} pad_1 & pad_2 & pad_3 \end{array} \\ \begin{array}{c} pad_1 \\ pad_2 \\ pad_3 \end{array} & \begin{pmatrix} 0.258 & 0.295 & 0.442 \\ 0.244 & 0.404 & 0.342 \\ 0.244 & 0.392 & 0.355 \end{pmatrix} \end{array}$$

$$P^4 = \begin{array}{c c} & \begin{array}{ccc} pad_1 & pad_2 & pad_3 \end{array} \\ \begin{array}{c} pad_1 \\ pad_2 \\ pad_3 \end{array} & \begin{pmatrix} 0.246 & 0.368 & 0.371 \\ 0.244 & 0.369 & 0.367 \\ 0.245 & 0.369 & 0.367 \end{pmatrix} \end{array}$$

It's All Just Matrix Theory?

$$P^5 = \begin{array}{c} \\ pad_1 \\ pad_2 \\ pad_3 \end{array} \begin{array}{c} \begin{array}{ccc} pad_1 & pad_2 & pad_3 \end{array} \\ \begin{pmatrix} 0.241 & 0.363 & 0.362 \\ 0.239 & 0.361 & 0.360 \\ 0.240 & 0.361 & 0.361 \end{pmatrix} \end{array}$$

$$P^6 = \begin{array}{c} \\ pad_1 \\ pad_2 \\ pad_3 \end{array} \begin{array}{c} \begin{array}{ccc} pad_1 & pad_2 & pad_3 \end{array} \\ \begin{pmatrix} 0.231 & 0.249 & 0.248 \\ 0.230 & 0.247 & 0.246 \\ 0.230 & 0.247 & 0.247 \end{pmatrix} \end{array}$$

$$P^7 = \begin{array}{c} \\ pad_1 \\ pad_2 \\ pad_3 \end{array} \begin{array}{c} \begin{array}{ccc} pad_1 & pad_2 & pad_3 \end{array} \\ \begin{pmatrix} 0.213 & 0.322 & 0.321 \\ 0.212 & 0.320 & 0.320 \\ 0.212 & 0.321 & 0.320 \end{pmatrix} \end{array}$$

It's All Just Matrix Theory?

- The probabilities of the three lily pads are .2, .3, and .3 no matter where the frog starts at the first step.
- the long range predictions are independent from the starting state
- the columns are about identical because the chain forgets the initial state
- but . . . this does not happen for all chains . . . .
  - we will show the conditions to be satisfied by the chain to guarantee this behaviour

It's All just Matrix Theory?

- consider π_0 is a vector of N component defining, $\forall$ state i, the probability that the Markov chain is initially at i

$$\pi_0(i) = \mathbb{P}\{X_0 = i\}, \text{ i=1 \ldots N}$$

- $\pi_n(j)$ defines the probability that the Markov chain is at j, after n steps.

$$\pi_n = \{\pi_n(1), \ldots, \pi_n(N)\}$$

Theorem

Let P be the transition matrix of a Markov chain, and let π_0 be the probability vector which represents the starting distribution. Then the probability that the chain is in state s_i after n steps is the ith entry of $\pi^n = \pi_0 P^n$

It's All Just Matrix Theory?

- $\pi_n(j)$ defines, $\forall$ state j, the probability that the Markov chain is at j, after n steps.

$$\pi_n = \{\pi_n(1), \ldots, \pi_n(N))$$

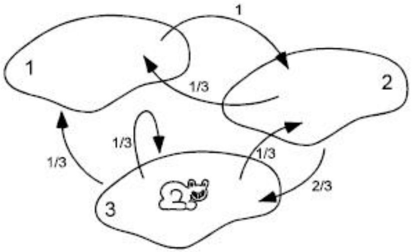
- thus, we have

$$\begin{aligned}\pi_1 &= \pi_0 P \\ \pi_2 &= \pi_1 P = \pi_0 P^2 \\ \pi_3 &= \pi_2 P = \pi_0 P^3\end{aligned}$$

- in general

$$\pi_n = \pi_0 P^n$$

It's All Just Matrix Theory?



- let us suppose $\pi_0 = (1/3, 1/3, 1/3)$
- we want to compute the state distribution after 3 frog jumps
- $\pi^3 = \pi_0 P^3 = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3}) \begin{pmatrix} 0.258 & 0.295 & 0.442 \\ 0.244 & 0.404 & 0.342 \\ 0.244 & 0.392 & 0.355 \end{pmatrix}$

It's All Just Matrix Theory?

No, not just matrix theory...

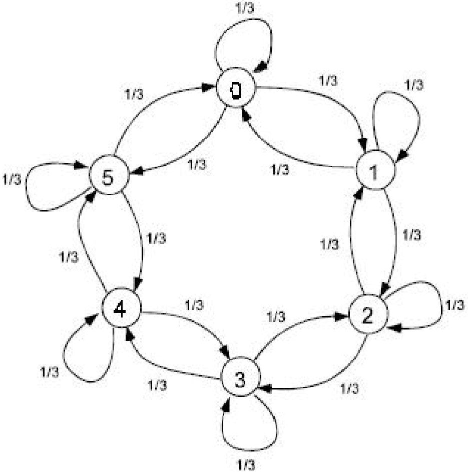
In principle, matrix algebra can give us the answer to any question about the probabilistic behaviour of a Markov chain, but this is not viable, in practice

Consider a P2P system

- Markov chains describe random walk on P2P overlays
  - states corresponds to peer
  - transition correspond to link overlays
- the size of the matrix is huge, matrix algebra may be not exploited

The same for social graph analysis

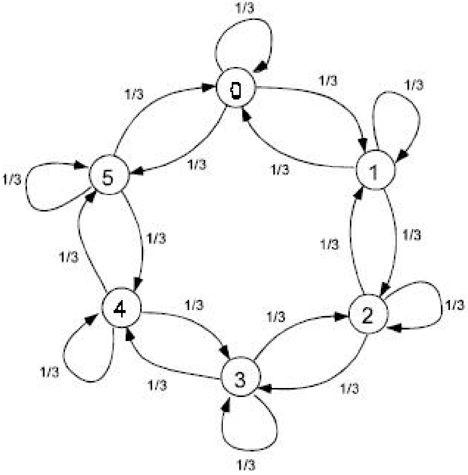
A Random Walk on a Clock



- a simplified clock with 6 number: 0,1,2,3,4,5
- from each state we can move clockwise, counter clockwise, or stay in place with the same probability
- the transition matrix is

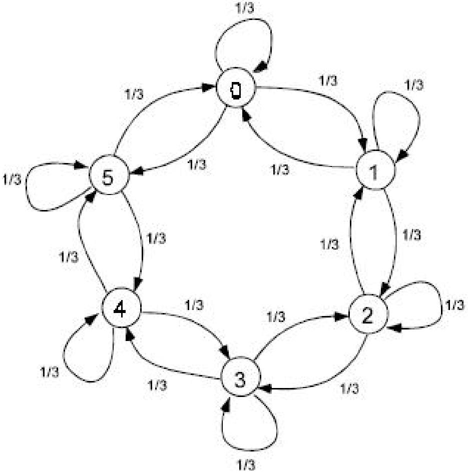
$$P(i,j) = \begin{cases} 1/3 & \text{if j=i-1 mod 6} \\ 1/3 & \text{if j=i} \\ 1/3 & \text{if j=i+1 mod 6} \end{cases}$$

A Random Walk on a Clock



- suppose we start out at $X_0 = 2$, that is $\pi_0 = (0, 0, 1, 0, 0, 0)$
- $\pi_1 = (0, 1/3, 1/3, 1/3, 0, 0)$
- $\pi_2 = (1/9, 2/9, 1/3, 2/9, 1/9, 0)$
- $\pi_3 = (3/27, 6/27, 7/27, 6/27, 3/27, 2/27$
- notice that the probability is spreading out away its initial concentration on the initial state 2

A Random Walk on a Clock



- now guess what is the state of th
random walk at time 10000
- an intuitive answer: X_{10000} is uniformly distributed over the 6 states
- this can be proven formally

$$\pi_n \to \left(\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}\right)$$

- after 10000 steps, the random walk has forgotten that it startec
out in state 2

Markov Chains: The Basic Theorem

Theorem (Basic Limit Theorem)

Any irreducible, aperiodic, Markov chain defined on a set of state S and with a stochastic transition matrix P has a unique stationary distribution π with all its component positive. Furthermore, let P^n be the n-th power of P, then $\lim_{n \to +\infty} P^n_{i,j} = \pi(j), \forall i, j \in S$

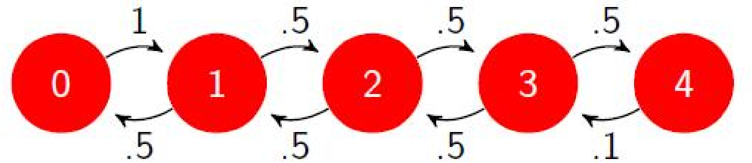
to understand the theorem, we need to define the words

- irreducible
- aperiodic
- stationary distribution

Before we shall see some examples where the limit does not exist

Random Walk with Reflecting Boundary(Periodic)

Periodic chain with period 2

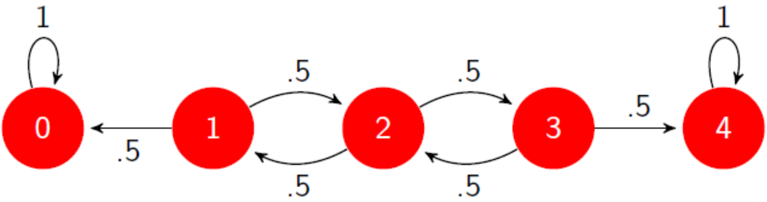


$$P = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0.5 & 0 & 0.5 & 0 & 0 \\ 0 & 0.5 & 0 & 0.5 & 0 \\ 0 & 0 & 0.5 & 0 & 0.5 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} \quad P^{2n} = \begin{pmatrix} 0.25 & 0 & 0 & 0.5 & 0.25 \\ 0 & 0.5 & 0 & 0.5 & 0 \\ 0.25 & 0 & 0.5 & 0 & 0.25 \\ 0 & 0.5 & 0 & 0.5 & 0 \\ 0.25 & 0 & 0.5 & 0 & 0.25 \end{pmatrix}$$

$$P^{2n+1} = \begin{pmatrix} 0 & 0.5 & 0 & 0.5 & 0 \\ 0.25 & 0 & 0.5 & 0 & 0.25 \\ 0 & 0.5 & 0 & 0.5 & 0 \\ .25 & 0 & 0.5 & 0 & 0.25 \\ 0 & 0.5 & 0 & 0.5 & 0 \end{pmatrix}$$

This means at every state i , if $P^n(i,i) > 0$, then n is even.

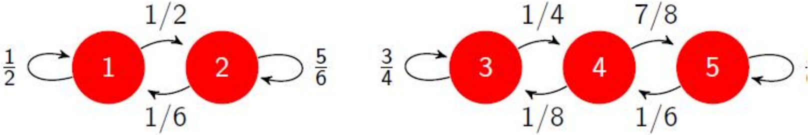
Random Walk With Absorbing Boundary



$$P = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.5 & 0 & 0.5 & 0 & 0 \\ 0 & 0.5 & 0 & 0.5 & 0 \\ 0 & 0 & 0.5 & 0 & 0.5 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} \quad P^n = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.75 & 0 & 0 & 0 & 0.25 \\ 0.5 & 0 & 0 & 0 & 0.5 \\ 0.25 & 0 & 0 & 0 & 0.75 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

- rows are not identical, i.e, the chain does not forget the initial state
- states 1, 2, 3 are transient: $p^n_{i,1} = p^n_{i,2} = p^n_{i,3} = 0 \ \forall i$, i.e. after passing a large amount of time the chain will stop visiting states 1, 2, 3.

Reducible Chains



$$P = \begin{pmatrix} 1/2 & 1/2 & 0 & 0 & 0 \\ 1/6 & 5/6 & 0 & 0 & 0 \\ 0 & 0 & 3/4 & 1/4 & 0 \\ 0 & 0 & 1/8 & 2/3 & 5/24 \\ 0 & 0 & 0 & 1/6 & 1/6 \end{pmatrix}$$

$$P^n = \begin{pmatrix} 0.25 & 0.75 & 0 & 0 & 0 \\ 0.25 & 0.75 & 0 & 0 & 0 \\ 0 & 0 & 0.182 & 0.364 & 0.455 \\ 0 & 0 & 0.182 & 0.364 & 0.455 \\ 0 & 0 & 0.182 & .364 & 0.455 \end{pmatrix}$$

Stationary Distribution

- a distribution which satisfies

$$\pi = \pi P$$

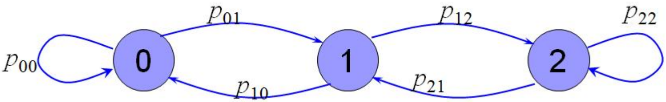
- every finite state Markov chain has at least one stationary distribution
- A trivial example with infinitely many stationary distribution: when P is the identity matrix, in which case all distributions are stationary
- The stationary distribution is the one obtained by computing the power of the transition matrix

Markov Chain Irreducibility

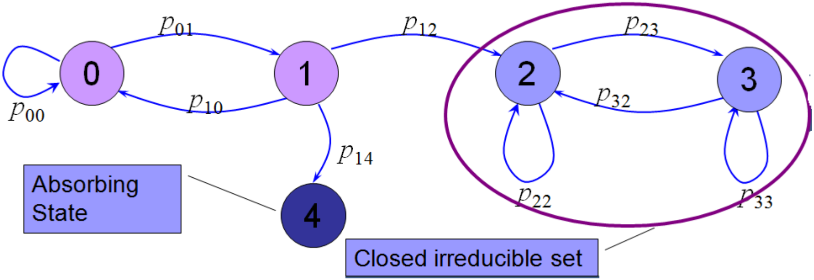
- State j is reachable from state i if probability to go from i to j in $n \geq 0$ steps is greater then 0
- A subset X of the state space S is *closed* if $p_{ij} = 0$, $\forall i \in S$ and $j \notin S$
- A closed set of states is *irreducible* if any state $j \in S$ is reachable from any state $i \in S$
- A Markov Chain is *irreducible* if the state space S is irreducible

Markov Chain Irreducibility

- Irreducible Markov Chain



- Reducible Markov Chain



Markov Chain Periodicity

Definition

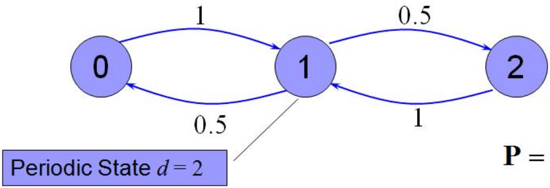
Given a Markov chain X_0,X_1,..., the period p_i of a state i is the greatest common divisor (gcd)

$$p_i = \{\text{gcd } n \text{ such that } P^n(i,i) > 0\}$$

Definition

An irreducible Markov chain is aperiodic if its period is 1, and periodic otherwise.

Markov Chain Periodicity



$$\mathbf{P} = \begin{bmatrix} 0 & 1 & 0 \\ 0.5 & 0 & 0.5 \\ 0 & 1 & 0 \end{bmatrix}$$

- If the period of state i is $\neq 1$, the chain returns to state i at regular times,
- If the period of state i is 1, the chain returns to state i can occur at irregular times
- A simple sufficient (but not necessary) condition for an irreducible chain to be aperiodic is that there exist a state i such that $P(i, i) > 0$

Markov Chain Recurrence

- *Hitting Time:* $T_{ij} = min\{k > 0 : X_0 = i, X_k = j\}$
- *Recurrence Time:* is the first time that the chain returns to state i
- Let ρ_i the probability that the chain will come back to state i, given it starts from state i
- A stete is *recurrent* if $\rho = 1$, otherwise it is *transient*.

Recurrent and Transient States